



## Building the Greater Philadelphia GeoHistory Network

**Saturday, December 3, 2005 Facilitated Session Transcripts**

### **Higher Education/Research –Table 1**

Our three big ideas:

First, it seems that in order for GIS to get moving in a big way we need to agree on some critical data that needs to be in digital form. Some of it may already be in digital form, census data and so on, that we just need to get from somebody else. But some kind of priorities seems to need to be made by a larger group to make sure that the kind of data that would make it easy to do very specialized research topics, because we plug it into maps, census data, immigration data, city directories and so on, that those critical base kinds of data are in existence.

Second, we must make sure that as much as possible the data is free and publicly accessible, and that it shouldn't require a lot of high tech, a lot of specialized money or access in order to get to things. We need to be able to make GIS data as available as possible.

Third, this is a kind of two-part issue. One part is in terms of audience, that the GIS raised the possibility of being able to plug into a larger regional or national system data that is gathered by local communities. People could do walking tours with a tape recorder that could just get loaded into neighborhood GIS information. At the same time, we're very much aware of two other problems. One is that there's a digital divide by class, by region, and so on, so that some areas could be developed much more easily than others. The other is a question of quality control. The more stuff that gets loaded in by a broader and broader range of people, the harder it is to actually keep control of what data we've got that's solid, that's authentic, so that people who plug into GIS systems could be confident that what they're looking at is reliable.

### **Higher Education/Research – Table 2**

Our group looked at a number of areas and finally settled on three main topics. The first is training, which includes promotion – there has to be word gotten out about GIS use as well as being able to provide the training, both help in house and online, to use GIS. The second is: it has to be user-friendly; it can't be so difficult that it requires a specialist to create or use. At the same time, it has to be secure. You want to be sure that the data

doesn't get corrupted, or hacked into – or that people aren't able to use the data improperly. The example for this is geological sites where people may go in and actually steal relics and other items. Finally, it has to be portable and accessible – if you can't move it or it takes too long to download because of the size of the data. One possibility is getting an Adobe PDF, where you can bring it in [inaudible]...on top of different other sites.

One of the things that came up after we finished the top three was the fact that we are talking about all these things you can use for GIS, but yet the basic premise is that we have a map on which you're plotting the points. In a city like Philadelphia you can plot fairly easily because you have a lot of data on Philadelphia. The further you go out in the suburbs – you're outside the cities into other areas; you don't have as much data and you need the maps to start with. Unless you have the points to link to, basically you're going to have difficulty trying to link materials to the points.

### **Higher Education/Research – Table 3**

Our group went way off in kind of left-brain directions, I'm afraid, we're going to get an F for having three things. The discussion was rich in invention and subsequent wrangling about how best to do this.

We could echo some of the things already said in our own ways about levels and kind of users. I remember David Seaman echoing others in the last couple of days, saying that everybody wants everything and they want it to be malleable – so mass and malleability are twin poles around which all of this revolves. We did talk about a system that might have a series of standard layers – the “GIS biggest hits for Philadelphia” – that would meet the needs of a lot of casual users, who are interested in browsing for certain kinds of information about the city; and a system that would allow for more sophisticated “I know what I want; I just want to get the data out of this and do my thing with it” kind of stuff. So that produced a lot of notes which, mercifully, no one will be able to read in a couple of days. We talked about keyword access, browsing access in this context; we talked about some basic conditions; we developed a list of things. One of the things we talked about too means for alerting others in the community that I have data – GIS blogging or some kind of clearinghouse for announcing the existence of data.

The most interesting thing we got into was with the final question on our list, a question that Walter Rice had set up yesterday about the system architecture. As we address these issues with GIS, we have to keep in mind that GIS is just another window into information that we're looking at here and these centralized clearinghouses really should just enable the PACSCL members to do what they already do best but do it for a larger audience. If a clearinghouse could be thought of as just a centralized dispatch system that would allow people to submit queries in free form, and then build enough intelligence in this system that it can go out and make queries to the individual repositories with whatever GIS spatial metadata being part of those queries, then that can happen in parallel with the member institution and only then has to be updated in that one location. I guess that was it, in a nutshell.

The other theme that floated through our conversation is about maps and their role in this. There's a lot of talk about boundaries as permanent and moveable. One of the things we talked about and I found very interesting was the idea "let's start with the data and see what kind of map gets built" rather than assuming that today's seventh ward is the idea we want to start with. So the idea of the map being the end product rather than the beginning of the search adds a dialectical twist to the discussion that we think bears a lot of examination, especially if users want to do something in addition to or other than finding out what my street looked like in 1850.

## **Business and Industry**

Our group was the business and industry group. Before I get to the big ideas, there were a few things that quickly emerged that were a little bit different from what perhaps the rest of you look at. I want to talk about those different assumptions that we ended up discussing around. One is that as businesses and vendors and service providers we are all both consumers and producers, both competitors and partners, and that many of the organizations we work for are often also in the same position and whatever we build has to be able to realize that those contradictory roles are going to be present by many of the users of such a system.

Much of the data that we work with doesn't belong to us. We are hired by all of you, much of the time, to do things with your data. Even if we wanted to make it available to the public, you often don't want us to, for very good reasons, whether they be privacy reasons or security around an archaeological site or liability reasons or public health reasons – the type of data that NIH has, for example, in aggregate.

Many of the cultural resource organizations we work with often do not have the money, technological capability to do some of the things or use some of the data that people have been talking about here, and probably never will. And yet those are precisely some of the organizations that we need to try to be supporting.

Those are the assumptions; those aren't the big ideas.

First. In many cases, there is an infinity of possible data and deciding which of those data are the most important may be very difficult. Therefore, we may not want to focus on data, because data is going to get produced, for a whole bunch of reasons that are driven by any number of reasons that will change over time. Rather, it may make sense to focus on services. What we're trying to construct is almost like a utility system, or electrical grid or water system. Thus, we want to focus on constructing services that would support a whole range of different data types.

What would be the low hanging fruit? One, that seems to come up a lot: people have got addresses, they've got place names, and they need to be able to geocode them. They don't need to be able to geocode them against what we usually think of as a geocoder in a GIS sense but in terms of the type of geocoder that Lex has been building for place names it would have to have both gazetteer-type data as well as historical street data as well as a

number of other aliasing systems that would have to be built in. So a geocode service seems like it would be enormously valuable, particularly for organizations that don't have a GIS, but simply have a list of addresses and they need to be able to find where they are so that they can use them in a geographic fashion. The second would be a type of search service, and I think others have discussed this as well so I won't elaborate on that. But the fundamental question, the scenario, is "I need to find all the stuff that's on the 800 block of Chestnut." That's the question that I think a lot of us face, certainly as vendors, but many of you do as well. And then the clearinghouse idea, the sense that this isn't a sort of master repository but it's a repository of links and pointers to where all of that information exists. So that those three services would be important.

The second big idea was that whatever you do, probably more important than the technology is that evangelism and outreach needs to be given a substantial budget in something like this. The type of thing that Maurie Kelly and the people at PASDA have done over the years to improve PASDA is not simply to add more data and document it and make the website better, but to drive around the state, year-in, year-out, visiting with people, teaching them about what they do, teaching them about metadata, offering to host the data there. That's a big effort; it's half of what PASDA does, and it probably should be here as well.

Finally, in many cases we don't know what everyone else is doing. A conference like this is very helpful, not because of what we talk about up there but because we have a chance to make connections and talk with each other and hear about other projects. That's not going to change. So, how do we find out about what other people are doing? Some people have suggested geoblogs; we were discussing a system we put in place in our office that will just handle knowledge within a very small group of people, and we're using a wiki. So what might a wikipedia for Philadelphia historic research look like? It wouldn't be like a blog, but it would be something that could be openly editable by a large group of people.

Some of these things are not digitizable. Some of it will be knowledge that is in people's heads. They're not willing to write it down, they're not going to put it in the wiki, they're not going to write about it on the blog. Events such as this, or even ones that are more focused on particular areas of shared interests, or even regular show-and-tells, like we all did in primary school, that kind of thing on a structured, regular fashion would be very valuable to fill in the gaps where the technology that we may deploy will never really replace the person-to-person contact that's very powerful in an event like this.

I want to close with one big question, and it's been raised by a lot of the different presentations and it's something that we face in our own offices. Oftentimes we talk about this information as if it's fairly static, as if the photo I have of the Christmas tree means the same thing now as it meant a hundred years ago. I love the diagram that Lex had – the shaft that was bending in one direction [ed. note: see Power Points from Merrick Lex Berman's Saturday presentation]. We have to realize that in the things that we do, their meanings to people change over time. Not only does our own sense of it change, but what it means to society and to culture changes and what we build needs to

accommodate that. These are not static things that we're documenting. They themselves change; as we learn more about things they create new complexes around them. So, does what we're going to build enable us to accelerate the process of incorporating that change over time? The image that one of the people at the table offered was "how does it go from my desk, where I'm a cultural resource service provider, to the state historical office, to a pocket culture browser? Faster?" Right now it probably doesn't – how do we accelerate that process and change and shifting understanding of what things mean?

### **Combined groups 5, 6, 7 – Government and publicly funded organizations**

We came up with three big ideas, and they reflect our nature as being publicly funded government institutions where we don't have to deal with a lot of the issues of copyright and other proprietary types of issues that some of your groups may have been struggling with.

Our three big ideas are that it should be cheap, easy, and standard. We need to be keeping in mind the needs of the diverse community that we're going to be servicing. It needs to be readily accessible to everyone. Some of the two different ideas that we talked about had to do with the fact that this could be more than just information about data. It could also serve as a window into existing collections, kind of a finding aid, helping folks identify where there might be sets of collections or sets of data that could be incorporated into a GIS system.

The other idea that we talked about had to do with the fact that it would be interesting to have the system have some contributory component, so when folks did some research they could basically come back and say I know about this data set or here's some information, and that's where the idea of standardization and quality control came into play. You wouldn't want anyone to mess with the integrity of your data, but you would want to be able to have people add to the richness of it.

#### **PACSCL**

c/o The Historical Society of Pennsylvania  
1300 Locust Street  
Philadelphia PA 19107  
215-985-1445  
<http://www.pacscl.org/>